Towards Generalist Biomedical AI

Year: 2023 | Citations: 386 | Authors: Tao Tu, Shekoofeh Azizi, Danny Driess, Mike Schaekermann, Mohamed Amin

---

Abstract

Medicine is inherently multimodal, with rich data modalities spanning text, imaging, genomics, and more. Generalist biomedical artificial intelligence (AI) systems that flexibly encode, integrate, and interpret this data at scale can potentially enable impactful applications ranging from scientific discovery to care delivery. To enable the development of these models, we first curate MultiMedBench, a new multimodal biomedical benchmark. MultiMedBench encompasses 14 diverse tasks such as medical question answering, mammography and dermatology image interpretation, radiology report generation and summarization, and genomic variant calling. We then introduce Med-PaLM Multimodal (Med-PaLM M), our proof of concept for a generalist biomedical AI system. Med-PaLM M is a large multimodal generative model that flexibly encodes and interprets biomedical data including clinical language, imaging, and genomics with the same set of model weights. Med-PaLM M reaches performance competitive with or exceeding the state of the art on all MultiMedBench tasks, often surpassing specialist models by a wide margin. We also report examples of zero-shot generalization to novel medical concepts and tasks, positive transfer learning across tasks, and emergent zero-shot medical reasoning. To further probe the capabilities and limitations of Med-PaLM M, we conduct a radiologist evaluation of model-generated (and human) chest X-ray reports and observe encouraging performance across model scales. In a side-by-side ranking on 246 retrospective chest X-rays, clinicians express a pairwise preference for Med-PaLM M reports over those produced by radiologists in up to 40.50% of cases, suggesting potential clinical utility. While considerable work is needed to validate these models in real-world use cases, our results represent a milestone towards the development of generalist biomedical AI systems.